

# Distribution of Maximal Luminosity of Galaxies in the Sloan Digital Sky Survey

E. Regős, A. Szalay, Z. Rácz, M. Taghizadeh and K. Ozogany

<sup>1</sup>CERN (Geneva), <sup>2</sup>Johns Hopkins U. (Baltimore), <sup>3</sup>ELTE (Budapest)

**Abstract.** Extreme value statistics (EVS) is applied to the pixelized distribution of galaxy luminosities in the Sloan Digital Sky Survey (SDSS). We analyze the DR8 Main Galaxy Sample (MGS) as well as the Luminous Red Galaxy Sample (LRGS). A non-parametric comparison of the EVS of the luminosities with the Fisher-Tippett-Gumbel distribution (limit distribution for independent variables distributed by the Press-Schechter law) indicates a good agreement provided uncertainties arising both from the finite size of the samples and from the sample size distribution are accounted for. This effectively rules out the possibility of having a finite maximum cutoff luminosity.

---

**Keywords.** Galaxies: formation, Galaxies: luminosity function, Cosmology: large-scale structure

## 1. Introduction

Extreme value statistics analyzes the behavior of the tails of distributions. The distribution of extreme values for i.i.d. (independent, identically distributed) variables converge to a few limiting distributions depending on the tail behavior of the parent population (Fisher-Tippet-Gumbell, Weibull, Fisher-Tippet-Frechet). The onset of the scaling behavior is quite slow, therefore requires very large samples. The galaxy samples in the SDSS redshift survey may be just large enough to attempt such an analysis, and we present here a study of the distribution of maximal luminosities.

We minimize the correlations between luminosities and positions by selecting the maximal luminosities from the batches of galaxies in elongated conical regions (defined by the footprints of HEALPix cells on the sky). We also explore random sampling.

The shape of the galaxy luminosity function is important for the EVS. This function is well described by the Schechter function, functionally similar (and motivated by) the theoretically derived Press-Schechter formula, with a power law distribution and an exponentially falling tail. Such a tail would imply a Fisher-Tippett-Gumbel (FTG) EVS distribution, with corrections for the finite sample sizes depending on the power law at low luminosities. In this analysis we will show that there is an excellent agreement with these expectations, implying that the Schechter function extends to very high luminosities, i.e. there is no indication for a sharp cutoff at a high but finite luminosity.

In order to arrive at the above results, it was important to notice that the galaxies can be divided into two types with significantly distinct luminosities and spatial distributions.

Previous studies have examined whether Brightest Cluster Galaxies (BCG) are the extremes of a red early-type galaxy luminosity function or they are different. Tremaine and Richstone (1977) investigated the statistics of the luminosity gap between the first and second brightest galaxies in clusters of galaxies and found that it is substantially larger than what can be explained with an exponentially decaying luminosity function. BCGs in high-luminosity clusters are not drawn from the luminosity distribution (LD) of all red cluster galaxies, while BCGs in less luminous clusters are the statistical extreme. For a volume-limited sample of Luminous Red Galaxies (many of them BCGs) we find that they are the extremes of a LD with an exponentially decaying tail.

Even though the SDSS sample is large, the residual from the FTG distribution can be explained only when we consider the corrections due to both the finite size of the samples and the distribution present in the sample sizes (the number of galaxies in a cone is finite and varies

from cone to cone). The distributions of the galaxy luminosities and the galaxy counts in pencil beams are constructed in Section 2.

## 2. Sample Creation

We use data from SDSS-DR8, available in a MS-SQL Server database that can be queried online via CasJobs (<http://casjobs.sdss.org>). The spectroscopic survey renders a complicated geometry defined by **sectors**.

We explore 2 different galaxy samples: the Luminous Red Galaxies (LRGs) (Eisenstein 2001) and the Main Galaxy Sample (MGS) (Strauss 2002). The LRGs (magnitude limit 19.2) are slowly evolving and intrinsically luminous red galaxies composed of old stellar populations, selected for tracing the structure at a higher redshift [0.20, 0.38] than MGSs. For selecting the MGS we impose a redshift interval of [0.065, 0.22] and magnitude [13.5, 17.65] in r-band Petrosian magnitude limit of 17.77. A total of  $N_T = 348975$  MGS and  $N_T = 52579$  LRG galaxies are obtained in a survey's comoving volume of  $V_S = 0.56 \text{ Gpc}^3$  and  $V_S = 2.18 \text{ Gpc}^3$  respectively.

### 2.1. Footprint and HEALPix based pencil beams, Distributions of galaxy counts in a HEALPix cell

In order to have close-to-i.i.d realizations (batches) from which to draw the maximal luminosities, we tessellate the sphere into regions defined by individual HEALPix cells (Gorski 2005), all of which have the same area. This creates 3-dimensional pencil-like beams that sample the galaxy populations across different redshifts.

We create the entire SDSS-DR8 spectroscopic footprint with resolution  $N_{side} = 512$  ( $\sqrt{\Omega_{pix}} \simeq 6.87'$ ). We degrade the footprint to 3 lower resolution maps defined by  $N_{side} = 16, 32$  and  $64$ , creating thus the cells that define the pencil beams. We use only the group of cells which satisfy that their fractional area occupancy inside the footprint  $f \geq 0.97$ .

The other approach we use is the random sampling of the luminosity parent distribution in batches of fixed size.

We construct the luminosity functions of the different galaxy samples and fit the MGS sample with Schechter, generalized gamma distribution (see also Figures 2 and 4).

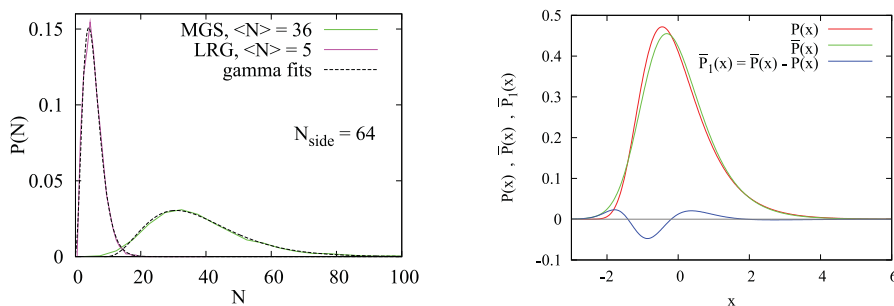
The LRG sample was made to include the brightest early types. Therefore, they are naturally fitted - their luminosity function itself - with an extreme value distribution or its special case the Gumbel.

## 3. Theory of Extreme Value Statistics

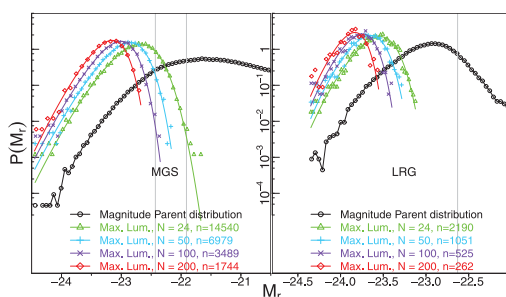
Extreme value statistics (EVS) is concerned with the probability of the largest value in a batch of  $N$  measurements. For us, they are galaxy luminosities in a given solid angle of the sky and  $N$  is the number of galaxies in the given angle.

The results of the EVS are simple for i.i.d. variables. The limit distribution belongs to one of three types and the determining factor is the large-argument tail of the parent distribution. Frechet type distribution emerges if the parent distribution  $f$  decays as a power law, Fisher-Tippett-Gumbel (FTG) distribution is generated by  $f$ s which decay faster than any power law and parent distributions with finite cutoff and power law behavior around the cutoff yield the Weibull distribution. All the above cases can be unified as a generalized EVS whose integrated distribution  $F_N(v)$  is given in the  $N \rightarrow \infty$  limit by  $F(v) = \exp[-(1 + \xi v)^{-1/\xi}]$  where  $\xi > 0, = 0, < 0$  correspond to the Frechet, FTG, Weibull classes, respectively, with the parameter  $\xi$  being the exponent of the power law behavior.

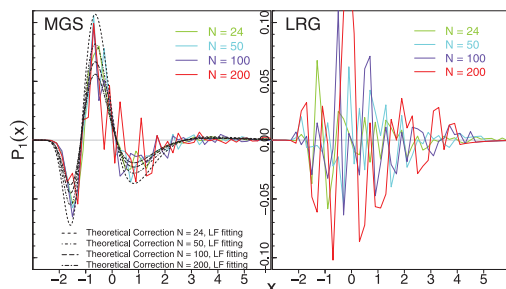
The parent distribution for galaxy luminosities is known, it is the Gamma-Schechter distribution. For this the limit distribution of extremal luminosities belongs to the FTG class ( $\xi \rightarrow 0$ )  $P(v) = \frac{dF(v)}{dv} = \frac{1}{b} \exp[-\frac{v-a}{b} - \exp(-\frac{v-a}{b})]$  where the parameters can be fixed by setting  $\langle v \rangle = 0$  and  $\sigma = \sqrt{\langle v^2 \rangle - \langle v \rangle^2} = 1$ . This choice ( $a = \pi/\sqrt{6}$  and  $b = \gamma_E \approx 0.577$ ) leads to a parameter-free comparison with the experiments provided the histogram of the maximal luminosities  $P(v)$  is plotted in terms of the variable  $x = (v - \langle v \rangle_N)/\sigma_N$  where  $\langle v \rangle_N$  is the average of



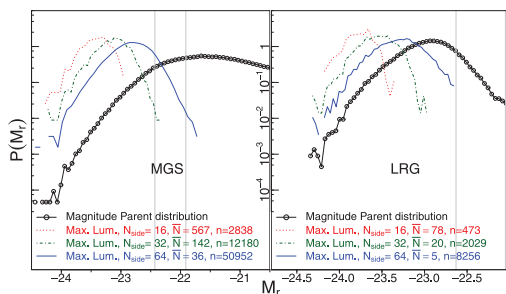
**Figure 1.** (a) Distribution of the number of galaxies  $N$  in the pencil beams for  $N_{\text{side}} = 64$ , for the MGS and the LRG samples. Fits to the empirical data are also shown. (b) Comparison of FTG with the count-averaged EVS, and their difference.



**Figure 2.** Distribution of r-band absolute magnitude  $M_r$  and distribution of maximal luminosity for the full samples according to the fixed batch size random sampling.



**Figure 3.** Empirical finite-size corrections from batches of fixed size  $N$  for random sampling.



**Figure 4.** Parent distribution of r-band absolute magnitude  $M_r$  and distribution of maximal luminosity for the full samples according to various HEALPix resolutions.

the maximal luminosity while  $\sigma_N = \sqrt{\langle v^2 \rangle_N - \langle v \rangle_N^2}$  is its standard deviation. The resulting scaling function is the universal function in the limit  $N \rightarrow \infty$ :  $\Phi_N(x) = \sigma_N P_N(\sigma_N x + \langle v \rangle_N) \rightarrow \Phi(x)$ .

### 3.1. Deviations from the ideal case

In addition to the assumption of i.i.d. variables, there are two additional problems with comparing data with theory. A notorious aspect of EVS is the slow convergence to the limit distribution. Second, the batch size  $N$  (the number of galaxies in a given solid angle) varies with the direction of the angle. Thus the histogram of the maximal luminosities  $P_N(v)$  is built from a distribution of  $N$ s. Both effects introduce corrections to the limit distribution.

### 3.2. Finite size corrections

Finite size corrections to first order can be written as  $\Phi_N(x) \approx \Phi(x) + q(N)\Phi_1(x)$ , where  $q(N \rightarrow \infty) \rightarrow 0$  and the shape correction  $\Phi_1(x)$  is universal. Both the amplitude  $q$  and the shape correction  $\Phi_1$  are known for Press-Schechter type parent distributions. The convergence is slow since  $q(N) = -\theta/\ln^2 N$ . The value of  $\theta$  is roughly 1 thus for characteristic range of  $N \approx 10 - 200$  one can expect a 20-4% deviation coming from finite-size effects.

The finite-size shape correction is  $\Phi_1(x) = [M_1(x)]'$  where  $M_1(x) = \Phi(x) \left[ \frac{ax^2}{2} - \frac{\zeta(3)x}{a^2} - \frac{a}{2} \right]$ .

### 3.3. Variable batch size

If the normalized distribution  $F(N)$  of  $N$  is known, the count-averaged EVS is

$$\bar{P}(v) = \lim_{N \rightarrow \infty} \int_0^\infty \mathcal{F}(N) P_N(v) dN \quad (3.1)$$

and the difference  $\bar{P}_1(x) \equiv P(x) - \bar{P}(x)$  provides us an estimate of corrections coming from the variable sample size.

We carried out simulations for the following:  $F(N)$  of the sample is fitted to  $(aN+b)^3 \exp(-(aN+b))$ , and  $F(N)$  is just the exact empirical distribution for the MGS and LRG samples (see Figures 1 and 3).

Combining the finite-size effects with the effects coming from the variable batch-size can produce the features observed in the deviations from the FTG limit.

## 4. Distribution of Maximal Luminosities, Discussion

Figure 2 shows the distribution of maximal luminosity for the MGS and LRG samples according to the fixed batch size random sampling. Figure 4 is the same but according to cones of various HEALPix resolutions.

The distributions obtained are FTG distributions. This effectively rules out the possibility of having a finite maximum cutoff luminosity.

We can claim that at large distances  $r$  the luminosity-luminosity correlations should decay with an exponent  $C(r) \sim r^{-\sigma}$  with  $\sigma > 2$ . The FTG-s found suggest weak correlations, and the noise and correlated features of the noise can be explained by the combination of finite-size and variable batch-size effects.

We find that the LRGs – many of them BCGs – are the extremes of a luminosity distribution with exponential decay.

## References

- Taghizadeh-Popp, M. *et al.* 2012, *ApJ*, 759, 100
- Fisher, R. A. & Tippet, L. H. C. 1928. *Proc. Camb. Phil. Soc.* 24,180
- Gnedenko, B. V. 1943. *Ann. Math.* 44, 423-453
- Górski, K. M. *et al.* 2005, *ApJ*, 622:759 , 109, 301